

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 25, 2002, 01:10:42 ; Search time 44 Seconds
(without alignments)
904.538 Million cell updates/sec

Title: US-09-708-724A
Perfect score: 2187
Sequence: 1 MGWVVVLCMKQLGAL.....LLAVTREGLEIRRIISRRAE 414
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Missing first 45 summaries

Database : PIR_73
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	138	6.3	1460	1 EDBE1F	immediate-early pr
2	128	5.9	2517	2 S58380	probable RNA-direc
3	121.5	5.9	1396	2 A44453	translation initia
4	120	5.8	1386	2 T49316	profilaggrin relat
5	117.5	5.4	1402	2 T46707	translation initia
6	116	5.3	606	2 T51880	hypothetical prote
7	116	5.3	3149	1 Q0BE8	BPLF1 protein - hu
8	115.5	5.3	1335	2 T18289	racGAP protein - s
9	114.5	5.2	501	2 S76563	hypothetical prote
10	114	5.2	1611	2 T38236	hypothetical prote
11	112.5	5.1	801	2 T29018	hypothetical prote
12	112	5.1	594	2 S33561	ref(2)p protein -
13	111	5.1	583	2 S29961	Ref(2)p protein -
14	111	5.1	728	2 S43768	transcription acti
15	110.5	5.1	743	2 C56695	transducin-like en
16	110.5	5.1	1012	2 T53172	period protein hom
17	110.5	5.1	1290	2 T00018	RAE-28 - mouse
18	110	5.0	817	2 S51342	verprolin - yeast
19	110	5.0	2142	2 T35098	MHC class III hist
20	109.5	5.0	590	2 S29964	ref(2)p protein -
21	109.5	5.0	1192	2 T18611	probable serine/th
22	109	5.0	883	2 S49126	brevican precursor
23	109	5.0	1246	2 G89287	protein H39E23.1 f
24	108.5	5.0	771	2 T38616	hypothetical prote
25	108.5	5.0	938	2 T39006	related to yeast z
26	108.5	5.0	1801	2 T26774	hypothetical prote
27	108	4.9	524	2 A75588	probable protein k
28	108	4.9	528	2 T47141	gastric mucin (clo
29	108	4.9	2187	2 T30826	nascent polypeptid

30 107.5 4.9 1208 2 T27822
31 107 4.9 1188 2 JC4889
32 107 4.9 1260 2 S60896
33 107 4.9 1388 2 A53317
34 106.5 4.9 365 1 SAVLWE
35 106.5 4.9 834 2 T42702
36 106.5 4.9 1074 2 T24877
37 106.5 4.9 1076 2 T24887
38 106.5 4.9 3262 2 AH2137
39 106.5 4.9 5262 2 T03454
40 106 4.8 620 2 S06733
41 106 4.8 633 2 S62057
42 106 4.8 805 2 T49385
43 106 4.8 2361 2 T25752
44 105.5 4.8 626 1 NBHUIA
45 105.5 4.8 1210 2 T39410

ALIGNMENTS

RESULT 1

EDBE1F

Immediate-early protein IE180 - suid herpesvirus 1 (strain Indiana-Funkhauser)
C:Species: suid herpesvirus 1
C:Date: 30-Jun-1990 #sequence_revision 30-Jun-1990 #text_change 16-Feb-1997
C:Accession: S04713
R:Cheung, A.K.

Nucleic Acids Res. 17, 4637-4646, 1989
A:Title: DNA nucleotide sequence analysis of the immediate-early gene of pseudorabies
A:Reference number: S04713; MUID:89315207; PMID:2546124
A:Accession: S04713
A:Molecule type: DNA

A:Residues: 1-1460 <CHE>

C:Superfamily: herpesvirus immediate-early protein IE175

C:Keywords: DNA binding; early protein; transcription regulation

Query Match 6.3%; Score 138; DB 1; Length 1460;
Best Local Similarity 22.8%; Pred. No. 0.21; Indels 104; Gaps 16;
Matches 76; Conservative 31; Mismatches 122;

Qy 76 LPTEEEFGLVQSMKCDTVRIKGVQLGP-----TTAPPLMTSEGNVTABDTEEA 124
Db 9 IETEGNFQLAAAAA-AAEEGIIASGPDGSGSGRRSGEDLLFGGGLFSDDAAEA 67
Qy 125 IRAFYVAVAA-----SAAEAHWRLVLLSGQIHEPIGSGGNIINTNGGRSCQNP 177
Db 68 EAAVLAAAGATRP RPSPSAQQQRHARR-----GSGEIVVLDDEDEDEEPG 114
Qy 178 LPSPDQSPSG--NATTSVTRDNYHLLTEEEFGVWSQMKWHSQKSG-GSVPVVRGPTQE 233
Db 115 SPAAG-SPVGLSIRAPSVT-----SSSGFGPGPAPGRR 149
Qy 234 PCSQSQILKESFVPPTTKENKQEREDENRLLPPPPVAETVPSPSVTEIETPLQIRP 293
Db 150 PROHSQ--RQRPGPPAAP-----GARPPPPPPPPPPPPA-----PPAPPAPR 192
Qy 294 TATTAGEPLGCTFTISPAFVHSLNKRQLELLREVEWPGRGHMAATCKQLQVEQD 353
Db 193 RPRGDGPPRGG-TRSVSPGRRG-LGPRRHQS-----QQRWFORRH----- 232
Qy 354 RTMSLAAAVREAPPPTGASSEPSVPALPGAD 386
Db 233 -----GGGFLQ-PPPPPGRRRRAAAAPPAE 259

RESULT 2

S58380

probable RNA-directed DNA polymerase (EC 2.7.7.49) - Crithidia fasciculata retrotrans
N:Alternate names: probable reverse transcriptase
C:Species: Crithidia fasciculata
C:Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 09-Sep-1997
C:Accession: S58380; S58381

R.Gabriel, A.
submitted to the EMBL Data Library, December 1994
A:Reference number: S58380
A:Accession: S58380
A:Molecule type: DNA
A:Residues: 1-2517 <GAB>
A:Cross-references: EMBL:U19151; NID:g624680; PID:g624681
R:Teng, S.C.; Wang, S.X.; Gabriel, A.
Nucleic Acids Res. 23, 2929-2936, 1995
A:Title: A new non-LTR retrotransposon provides evidence for multiple distinct site-specific recombination events in the human genome
A:Reference number: S58381; MUID:9388509; PMID:7659515
A:Accession: S58381
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1825-2116 <TEN>
A:Cross-references: EMBL:U19151
A:Experimental source: retrotransposon CRE2
C:Genetics:
A:Mobile element: retrotransposon CRE2
C:Keywords: nucleotidyltransferase

Query Match 5.9%; Score 128; DB 2; Length 2517;
Best Local Similarity 21.4%; Pred. No. 2;
Matches 90; Conservative 41; Mismatches 156; Indels 134; Gaps 19;

QY 5 SVVVLVCGMKQLGALQASVLSIITENQKRCPCGAGNLTQCNLTLPVSVSHRSPGPN 64
DB 1123 SIVDVAIRMKKEISTTDLTAALGIRKQTSDACPVCDAMQSTIQH-----VPGK 1172
QY 65 AAV-----SVTGGDCHLPTTEEEFGVLVOSMKCDVIRKGVLOGPTTAPPLMTSGNVT 118
DB 1173 AVNFSEIEATATGNQI-----TTVV-----APPTTFHK-----A 1201
QY 119 EDTEAIRAFVAVAAASAAEAWHRLVLLSGQTHEPIGSGNIIN-----TNKGGRC 173
DB 1202 PNGAEVTRSLLAVIDTTGPEA-----QVHRVTANGTAQAKCKVLPGEGGE 1250
QY 174 QNPALPSPQPSGN-----ATTSVTRDNYHLLTTEEEFGVMSQSMKWHSONKS 221
DB 1251 TRKPTVP-----PKGNLLCYLHKPRVTQTAQRTEDDDHRTTNE-EISSSSHOAPKKR 1304
QY 222 GGSVPVGRGTQEP-CSESQILKESFVPTTP-----KENNKQEREDENWRLPPPPVAET 274
DB 1305 PASSTTTPSMKPCNSCEVKI-----DTTGRQRPNNREEQTRTARAPTYTDNA 1357
QY 275 PVPSVTEIETPQRIPTATIAEPLGCHCTTISPAFVHSLNKRKQLELLREVEW 334
DB 1358 ANENASARKREE-QR-----GLGGATST-SPVINS-----SADFSWVEDTAT 1400
QY 335 PGRGHMAATCCKLQVEGQDRTMSLAAP-----VREAPPPPTGASSEP 377
DB 1401 PIQRTMSFTLDPLSPDEDE-----AVAPDPFLASIVQVVEEDEGTAEAPSMNGENSEP 1455
QY 378 S 378
DB 1456 A 1456

RESULT 3

A44453
translation initiation factor eIF-4 gamma - human
C:Species: Homo sapiens (man)
C:Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 08-Dec-1994
A:Title: Amino acid sequence of the human protein synthesis initiation factor eIF-4 gamma
A:Reference number: A44453; MUID:93054654; PMID:1429670
A:Accession: A44453
R:Yan, R.; Rychlik, W.; Etchison, D.; Rhoads, R.E.
J. Biol. Chem. 267, 23226-23231, 1992
A:Title: Amino acid sequence of the human protein synthesis initiation factor eIF-4 gamma
A:Reference number: A44453
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-1396 <YAN>
A:Note: sequence inconsistent with the nucleotide translation

A:Note: sequence extracted from NCBI backbone (NCBIP:118286)

Query Match 5.6%; Score 121.5; DB 2; Length 1396;
Best Local Similarity 24.1%; Pred. No. 2.6;
Matches 60; Conservative 28; Mismatches 92; Indels 69; Gaps 11;

QY 167 NKGGRSCONPALPSPDQSPSGNATTSVTRDNYHLLTTEEEFGVMSQSMKWHSONKSGGVP 226
DB 40 SQGAIIADRPGLPGGHEHSPS-----ESQPSSPSPPT 70
QY 227 VRGPTQEPCESEQILKESFVPT-----TPKENNKQERE-DENWRLPPPPVAETPVSP 279
DB 71 SPSPVLEPGSEPNLAVLSIPGDTMTTIQMSVEESTPISETGEYRLSP-----TPLAEP 127
QY 280 SVTEIETPLQRIPTATIAEPLGCHCTTISPAFVHSLNKRKQLELLREVEWPGRGH 339
DB 128 -ILEVEVTLKRPVESEFSSSPLQAPT-----PLASHTV-----EIHEP-NGM 168
QY 340 MAATCCKLQVEGQDRTMSLAAPVREAPP-----PT-----GASSEPSVPALPGADPQR 389
DB 169 VPSEDEPEVESSPELAPPAC-SESPVPIAPTQAPPELLNGAPSPPAVDLSPVSEPEE 227
QY 390 SAEILLAV 398
DB 228 QAKEVTASV 236

RESULT 4

T49316
profilaggrin related protein [Imported] - Neurospora crassa (fragment)

N:Alternate names: protein B13N20.10
C:Species: Neurospora crassa
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T49316
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatu
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022
A:Accession: T49316
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1386 <SCH>
A:Cross-references: EMBL:AL355925; GSPDB:GN00116; NCSP:B13N20.10
A:Experimental source: BAC clone B13N20; strain OR74A
C:Genetics:
A:Gene: NCSP:B13N20.10
A:Map position: 6

Query Match 5.5%; Score 120; DB 2; Length 1386;
Best Local Similarity 18.6%; Pred. No. 3.3;
Matches 94; Conservative 66; Mismatches 144; Indels 202; Gaps 23;

QY 53 LPSVSHRSPPGNAAY-----SVTGG-----DCHLPTTEE----- 81
DB 637 IPEISLQGPENASLRPGPEVLDROSQANGKVSQSGALSERSQLPTQGEAEKLGTTT 696
QY 82 FGVLVQSMKCDT-VRIKGVLOGPTTAPPLMTSEGNVTAEDEEAFVAVAAASAAEA 140
DB 697 IGRVPSQRLSVPSQKPPQPSPLGTAPVLQSSSRPLQPAQVRLHSSSIIGSSSPQ- 755
QY 141 WHWRLVLLSGQIHEPIGSGN-----IINTNKGGRSCONPALPSPQSPSG 187
DB 756 -----SGSIYTPVQKQNFVRRPGAKLPLFERDGGTKGTVTALPLK9HSSG 806
QY 188 NATT--SVTRDNYHLLTTEEEFGVW-----SQSMKWHSONKSGSVVVRGPTQEP 234
DB 807 POTTNSITDRPKQLL-----FGLKESIKLFSDDPPAKSVVEYGENEENEE---QSPEDM 859
QY 235 CSQSQIL-----KESFVPP-----TTPKENNKQEREDENWRLPPPPVAETPV- 278
DB 860 DDYDEIVELEEEQSSAESEFEPPLRLSLPDKNSSPIKS-----ALRPPTKGHTPGRAVQFA 915
QY 279 PSVTEIETPQRIPTATIAEPLGCHCTTISPAF-----VHSLNKRK----- 322

Db 916 NSVRSPTRRNADVEORRPFVGGRTPTVIITSPSTHTSSVHPAGSBERIEENATCM 975
 Qy 323 -----ROLEL-----LLREVEW-----PGRG 338
 Db 976 YEEDYQDEQEHELDVTDMDDMDVDDDIETWEHSFAPAPAVQAPRPPRPLSPGRG 1035
 Qy 339 HMAATCCCKLQVEGQDRTMSLAAPVREAP----- 367
 Db 1036 RQ-----RQOQQOQELQOQVAAAP-RQPPAAQPLIKTSTLENSPVIRSGVSLFGYGR 1089
 Qy 368 ---ppptGASS-----EPSPALP 383
 Db 1090 KSLPPPIPPSNRNIVTQLQPLPSLP 1115
 RESULT 5
 146707
 translation initiation factor eIF4-gamma - rabbit
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 14-Feb-1997 #sequence_revision 14-Feb-1997 #text_change 31-Mar-2000
 C:Accession: I46707
 R:Yan, R.; He, W.; Rhoads, R.E.
 J. Biol. Chem. 268, 19200-19203, 1993
 A:Title: Mapping the cleavage site in protein synthesis initiation factor eIF-4 gamma of
 A:Reference number: I46707; MUID:93374895; PMID:8396129
 A:Accession: I46707
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1402 <YAN>
 A:Cross-references: GB:I2202090; NID:g404774; PID:AAA31242.1; PID:g404775
 Query Match 5.4%; Score 117.5; DB 2; Length 1402;
 Best Local Similarity 26.4%; Pred. No. 5;
 Matches 63; Conservative 18; Mismatches 89; Indels 69; Gaps 11;
 Qy 159 SGGNIINTNKGSRCONPALSPDQSGNATTSVTRNYHLLTTEEFVWSQSMKWSH 218
 Db 40 SOGAI-----GGR-----PGLGPEHSPS-----ESQ 62
 Qy 219 NKSGSVVRGPTQPCSESOILKESFYPTTPKENNKQEREDENWRLPPPPVATPTVP- 277
 Db 63 PSSPSPTSPPILEFGSEPNLAVLS-LPGDTMTSGMTQMPVEE-----PAPISREAGEPY 117
 Qy 278 ---SPSVTEIETPLQRIPTATIAAGEPLGHCTFTISPAFVSHVNLNKRKQLELLREVEWP 335
 Db 118 CLSPETPLAEIPEVETLS---KVPVSESSPIQVLTPLASHK-----MEIHEP 167
 Qy 336 GRGHMAATCCCKLQVEGQDRTMSLAAPVREAPPPPTGASSEPSVPALPGADPQRSALL 394
 Db 168 NG-----VPSSEDELEVESSP-ELAPPPPPACLSSESPVPIAPTQPE-----ELL 212
 RESULT 6
 T51880
 hypothetical protein DKFzp762G2015.1 - human (fragment)
 C:Species: Homo sapiens (man)
 C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 20-Oct-2000
 C:Accession: T51880
 R:Ansorge, W.; Wilkner, U.; Mewes, H.W.; Weil, B.; Wiemann, S.
 submitted to the Protein Sequence Database, July 2000
 A:Reference number: 225852
 A:Accession: T51880
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-606 <AAA>
 A:Cross-references: EMBL:AL390166
 A:Experimental source: adult melanoma (Mewo cell line); clone DKFzp762G2015
 C:Genetics:
 A:Note: DKFzp762G2015.1
 Query Match 5.3%; Score 116; DB 2; Length 606;
 Best Local Similarity 22.2%; Pred. No. 2.2;
 Matches 102; Conservative 38; Mismatches 156; Indels 164; Gaps 22;

Qy 10 VCGMKOLQALQASVSLSIITENQCKRCPFCGAQNLMTCONPTLPSVSHRPPGNAAVSV 69
 Db 120 VGLP-CARMTTDS-----HKASPEDGQGL---PQPKRPLADGSGPGPEGK 166
 Qy 70 TGGDOCHLPTTEFEGVYVQSMKCDTVRIKGVLO-----GPTTAPPLMTSEGNVTAEDTE 122
 Db 167 VGLLNHRPVAMDAG-----DSADQSGERKDKESPRAGPT--EPMDTSEATVCHSDLE 216
 Qy 123 EAIRAFVYVAAAASAAEAHWHRLVLLSQIHEPIGSGNININKNKGRSCQ----- 174
 Db 217 RT-----PPLLGCRPARNRGPERSPRTSELSLELSI 246
 Qy 175 -----NPALSPDQSPSGNATTSVTRDNYH-----LTTEEFG----- 207
 Db 247 SARQOFTPLTQAQAPAPAP---ATTCTRAGGHPPEPLSLSRKRKLLDETSKTL 303
 Qy 208 ----VWSQSMKWSHQNKGSGVVRGPTQPCSES-----QILKESFY----- 246
 Db 304 DAYRVWQOGQK-----GVAYDLGRVERIMSETYMLIKQVDEEALEQAQVKFCQVHLG 355
 Qy 247 -----PTTTPKENNKQEREDENWRLPPPPVATPTVPSPSVTEIETPLQRIPTATI 297
 Db 356 LPPDRDRPTTTPK-HPKDQREN-FFPVTVVPTAPDPADSV-ORPSDAHTKRPALA 412
 Qy 298 AGEPLGHCTFTISPAFVSHVNLNKRKQLELLREVEWPGRGH---MAATCCKLQVEG--- 351
 Db 413 AATTI-----ITCPSASASTLDQSKD-----SGPRPRPRPEATPSMASLGP 459
 Qy 352 ----QDRTMSLAAPVREAPPPPTGASSEPSVPALPGADP 387
 Db 460 PEPRRDEGAQEAASETQPLSSPPTAASK--APSSGSAQ 497
 RESULT 7
 Q00B8
 BPLF1 protein - human herpesvirus 4 (strain B95-8)
 C:Species: human herpesvirus 4, Epstein-Barr virus
 C:Date: 25-Feb-1985 #sequence_revision 25-Feb-1985 #text_change 16-Jul-1999
 C:Accession: G93085; A03747; S32993
 R:Bankier, A.T.; Deininger, P.L.; Farrell, P.J.; Barrell, B.G.
 Mol. Biol. Med. 1, 21-45, 1983
 A:Title: Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8 Epstein-Barr vi
 A:Reference number: A93085; MUID:85035713; PMID:6092825
 A:Accession: G93065
 A:Molecule type: DNA
 A:Residues: 1-3149 <BAN>
 A:Cross-references: EMBL:V01555; NID:g59074; PID:CAA24839.1; PID:g1334853
 R:Baer, R.; Bankier, A.T.; Biggin, M.D.; Deininger, P.L.; Farrell, P.J.; Gibson, T.J.
 Nature 310, 207-211, 1984
 A:Title: DNA sequence and expression of the B95-8 Epstein-Barr virus genome.
 A:Reference number: A03794; MUID:84270667; PMID:6087149
 A:Contents: annotation; protein coding region
 C:Superfamily: human herpesvirus 4 BPLF1 protein
 Query Match 5.3%; Score 116; DB 1; Length 3149;
 Best Local Similarity 24.6%; Pred. No. 17;
 Matches 81; Conservative 27; Mismatches 111; Indels 110; Gaps 18;
 Qy 103 PTTAPPLMTSEGNVTAEDTERAIRAFVYVAAAASAAEAHWHRLVLLSQIHEPIGSGGN 162
 Db 340 PASAAPASAPASAPASAPASAPASAPASAPASAPASAPASAPASAPASAPASAP 390
 Qy 163 LINTNKGSRCONPALSPDQSP---SGNATTSVTRDNYHLLTTEEFVWSQSMKWSHQN 219
 Db 391 -----TPGVPAFTPPRASSGAAPQTPKARKK-----GLGKDSPP-HKK 427
 Qy 220 KSGSVVRGPTQPCSESOILKESFYPTTPKENNKQEREDENWRLPP-----PP 270
 Db 428 TSGRRLPLSSITD---TEDDQLPRTHVPPHPPPSAA-----RLPPPIPIPHQSP 475
 Qy 271 VAETVP-----SPSVT-----EITPL-----QRIPTA-----TIAGEPLGHC 305

Db 476 ASPTHAPDVSTIAPSVTPSKPLQIPIPLPQAAAPSNPKIPLTTPSPSPATAAAP---T 532
QY 306 TTTTSPAFVHSLVNLKRRKQLELLREVEWPGRGHMAATCKKLQVEGQDRTMSLAAAP--- 362
Db 533 TTTTSPPTTQ-----QQPPQS--AAPAPSPLLPQQQPTPSAAAPAPSP 573
QY 363 VREAPPPPTGASSEPSVP-----ALPGADP 387
Db 574 LPQQQPPPSAARAPSPLPQQQPLPSATP 602

RESULT 8
T18289
racGAP protein - slime mold (Dictyostelium discoideum)
C:Species: Dictyostelium discoideum
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jul-2000
C:Accession: T18289
R:Judson, S.B.; Eccleston, J.; Strom, M.
J. Biol. Chem. 272, 15682-15685, 1997
A:Title: Cloning of a rhoGAP homolog from Dictyostelium discoideum.
A:Reference number: Z18858; MUID:97332648; PMID:9188459
A:Accession: T18289
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1335 <LUD>
A:Cross-references: EMBL:Y10159; NID:g2190354; PIDN:CAA71241.1; PID:g2190355
C:Genetics:
A:Gene: racGAP
A:Introns: 113/1; 205/3

Query Match 5.3%; Score 115.5; DB 2; Length 1335;
Best Local Similarity 20.3%; Pred. No. 6.4;
Matches 69; Conservative 49; Mismatches 155; Indels 67; Gaps 12;

QY 16 LQALQASVLSLIITENQKRCPCGACQ-NLMTQNPLPVSVSHRSPGCGNAAVSVTTGGDC 74
Db 179 LNQIMEDSAVYTSIIKTVEFISFKOSNHLSTIDIPISVSLPKSPNSVSVS-SGGDS 237

QY 75 -----HLPTTEEFGLVQSMKCDTVRIKVLQGP----- 103
Db 238 VSSLBEENLYATALPYQASQGWHLFPKDDKXIVLIDIKSEGWLKGELNGKIGTFPASY 297

QY 104 -----TTAPPLMTSEGNVTAE--DTEAIRAFVYAVAAASAPAHWRHLVLLSGOIHPEI 157
Db 298 VEIIAIPPVGVPTQLIPDPDSQDLSTISSPLSSSTSS-----SSSISTDS 348

QY 158 GSGGNIINTKGRSQCNALPSPQSPSGNATTSVTRDNYHLLTEEEFGVWSQS----- 212
Db 349 NLSSNNNNNNNNNNSTPILSS--TSTTTTTTTTNNNNNNNTFQPIVLSKSSFSKS 406

QY 213 -MKWHSQKSGSVVRGPTQPCSESIQKESFVPTTPKENNKOEREDENWRLPPPPV 271
Db 407 TISTNPSSKSSNLLISNP--PPVKNSTTATSSPDIT-----KLIIGARNKLIAPPL 459

QY 272 AETPVPSVTEIEPLQRIPTATAGEPLGHCTFTTSP 311
Db 460 VPPTETAPPPLSAPPPL--ISRTSTLTSS--SNTAFTAPP 495

RESULT 9
S76563
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C:Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S76563
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.; O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis s.
A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76563

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-501 <KAN>
A:Cross-references: EMBL:D64002; GB:AB001339; NID:gl001612; PIDN:BAAL0409.1; PID:gl01
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 5.2%; Score 114.5; DB 2; Length 501;
Best Local Similarity 23.7%; Pred. No. 2.2;
Matches 94; Conservative 42; Mismatches 121; Indels 139; Gaps 24;

QY 36 RCPFCGAQLMTCQNPTLPSV-----SHRS---PPGNAASVVTGGDCHL--PTEEEFGVL 85
Db 163 RLPITG-----PPIPSLPRFSTRIQIPRPPGLFAPTPDQGMNPLGPGNRPDFGL 213

QY 86 VQSMKCDTVRIKVLQGPPT-APPLMTSEGNVTADTE-EAIRAFVYAVAAASAAEAHW 143
Db 214 PPSQ--GTGFTIRAPQGITPQEPGLRAEG-----EDTSPEKLENT-LVSIQAQOQALKA 266

QY 144 RHLVLLSGOIHPEIGSGGNIINTKGRSQCN--PALSPQSPSGNATTSVTRDNYHLL 201
Db 267 KDTI--SG-AVPPI-----ACRNRTEATVVYVNFPSGQK-----DIV 300

QY 202 TEEERGVWSQSMKHSQKSGSVVRGPTQPCPS---ESQIL--KESFVPTTTPKENNK 256
Db 301 GRSYPIIFNQLAQIAIARTYGE-----PTQVTVSFKYDAEICGGVDQFIPB----- 347

QY 257 OBEREDENWRLP--PPVVAETPV---PSPSVTEIETPLQRIIP-----RTATIAEGLGH 304
Db 348 -GGETTNPVAVTVPVSQVTPAPTISPAPGIAPSPAPLQPTPPPAVRSRSPMPDAPAPR 406

QY 305 CTTFTISPAFVHSLVNLKRRKQLELLREVEWPGRGHMAATCCKLQVEGQDRTMSLAAAPVR 364
Db 407 RQPTTTPS-----DPPMNVAPSPTR 426

QY 365 EAPPP-----PTGASSEPSVPALPCADPQRSAILLL 395
Db 427 SAPAPAPTATPTTSPQSLPKTKG-----EMLL 455

RESULT 10
T38236
hypothetical protein SPAC23A1.17 - fission yeast (Schizosaccharomyces pombe)
C:Species: Schizosaccharomyces pombe
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C:Accession: T38236
R:Murphy, L.; Harris, D.; Wood, V.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, February 1998
A:Reference number: Z21780
A:Accession: T38236
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1611 <MUR>
A:Cross-references: EMBL:AL021813; PIDN:CAA15991.1; GSPDB:GN00066; SPDB:SPAC23A1.17
A:Experimental source: strain 972h; cosmid c23A1
C:Genetics:
A:Gene: SPDB:SPAC23A1.17
A:Map position: 1

Query Match 5.2%; Score 114; DB 2; Length 1611;
Best Local Similarity 22.4%; Pred. No. 10;
Matches 96; Conservative 35; Mismatches 168; Indels 130; Gaps 19;

QY 17 GOALQASVLSIITENQKRCPCGACQNLMTQNPTLPSV-----SVSHRSP-PCGAAV 67
Db 804 GSQLNRVVPSIVT--SSGR-----FALPDEMASPSSSICHPLPSPPPADF 847

QY 68 SVTGGDCHLP-----TEEEFGLVQSMKCDTVRIKVLQGPPTAPPLMTS 112
Db 848 NSLNVDFEPEHSYLSPEAPQPSVEEESFNAT-----VIHAPT--PSTATF 892

QY 113 EGNVT-----AEDTEAIRAFVYAVAAASAAEAHWHLVLLSGOIHPEIGSGNI 163
Db 893 QGHPTISNVATPPLKQDVTESKASPVADASATHQSSTGLTQITQLGSMRLPT-----KL 948


```
Db 77 KCESNNHVOAFLAPVEEPKATQKQSSANAEPVDDPSNFTTIHDAVEDCDGGLAPLIG 136
QY 128 FYVAASAAAEAW-----HWRLVLLSGIHEPIGSGNIIIN---TNKGGRSQNP 176
Db 137 FRYKVCQSNYDLCKKCELAHKHPELML-----RMPTNNGGWDVDAWFTGGRSRGHCP 191
QY 177 ALPSPDQSPGNATTSVTRDNYHLITTEEFVWQSMKWHQSO---NKSGSVFVR----- 228
Db 192 FQETNQADPAGEPARDSRRERQ--ARRHAGVLSQFVEMMTNPLNLTATTATAPAEPPQPK 249
QY 229 -----GTPQPCSESIQLKESFVPPPTPKENNKQEREDENWRLLPPVAETPPVPSV 281
Db 250 AAEQTESPPQAEPTVTAKEASEAKTEPKKVNTDQ-----SVPTEDPVTTPRS 300
QY 282 TELETPLQRIPTANTAGEPLGCHTFTTISPAFVHS---VLNKKRKQLELLREVEWPGRG 338
Db 301 TQPTTPVINDNISQI-----VPEYMRAGIELLNFSEMFSLIITTEGGDSG 349
QY 339 HMAATCC-----KLQVEGQDRMTSLAAAPVREAPPPPTGASSEPSVPALPGADPORSAB 392
Db 350 IFAPSTPSAENKPKPEQSQSQSGQSSANQSAVPSAAPSANQATPSIGSIDPAQLE 409
RESULT 14
S43768
transcription activator Vp1 - rice
C:Species: Oryza sativa (rice)
C:Date: 10-Dec-1994 #sequence_revision 26-May-1995 #text_change 20-Jun-2000
C:Accession: S43768; S48899
R:Hattori, T.; Terada, T.; Hamasuna, S.T.
Plant Mol. Biol. 24, 805-810, 1994
A:Title: Sequence and functional analyses of the rice gene homologous to the maize Vp1.
A:Reference number: S43768; MUID:94250843; PMID:8193305
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-728 <HAW>
A:Cross-references: EMBL:D16640
A:Experimental source: strain Nipponbare
R:Hattori, T.; Terada, T.; Hamasuna, S.
submitted to the EMBL Data Library, July 1993
A:Description: Sequence and functional analysis of the rice gene homologous to the maize
A:Reference number: S48899
A:Accession: S48899
A:Molecule type: DNA
A:Residues: 1-670, 'p', 672-683, 'p', 685-728 <HAW>
A:Cross-references: EMBL:D16640; NID:g391884; PIDN:BAA04066.1; PID:g391885
A:Experimental source: strain Nipponbare
C:Genetics:
A:Gene: Vp1
A:Introns: 526/3; 556/3; 590/2; 606/1; 631/3
C:Superfamily: rice transcription factor Vp1
C:Keywords: DNA binding; transcription factor
Query Match 5.1%; Score 111; DB 2; Length 728;
Best Local Similarity 22.2%; Pred. No. 6.1;
Matches 67; Conservative 29; Mismatches 114; Indels 92; Gaps 15;
QY 170 GRSCONPALSPD-----QSPSGN-----ATSVTRDNYHLITTEEFVWQSMKWHQSO 218
Db 257 GAASDPQPLPSGANPGYEFPSGGEMGSAATS-----WMPYQAFTPP 300
QY 219 NKSGSVFVRG-----PTQPCSESIQLKES---FVPTTPPKENNKQEREDENWRLLP----- 267
Db 301 AAYGGDAMYPGAAGFPFQSCSKSVVSSQFPFPTAAAAGDMHASGGNWNAPQFPA 360
QY 268 PPVPAETPVPSVTEIETPLQRIPTANTAGE-PLGHCTFT-----ISPAFVHSLYNK- 320
Db 361 PFPVST--SSYTMFSVVP-----PFTAGFPGQYSGGHAMCSPRLAGVPSPTKEARKKR 414
QY 321 --RKQLELLREVEWPGRGHMAATCCCKLQVEGQDRMTSLAA----- 361
Db 415 MARQRLSLCQQO-----RSQOLNLSQIHISGHPOEPSRAAHSAVPYTPSSAGCRSWGI 468
```

```
QY 362 -----PVREAPPPTTGCASSEPS-VPALPGADPORSABELLLAVTREGLEIRIISR 410
Db 469 WPPAAQIQIIONPLSNPNPPATSKQPKSPKPKPQAAA-----TAGAESLQSTASE 523
QY 411 KR 412
Db 524 KR 525
RESULT 15
C56695
transducin-like enhancer-of-split homolog TLE-2 - human
C:Species: Homo sapiens (man)
C:Date: 21-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 26-May-2000
C:Accession: C56695
R:Stifani, S.; Blaumueller, C.M.; Redhead, N.J.; Hill, R.E.; Artavanis-Tsakonas, S.
Nature Genet. 2, 119-127, 1992
A:Title: Human homologs of a Drosophila enhancer of split gene product define a novel
A:Reference number: A56695; MUID:93265135; PMID:1303260
A:Accession: C56695
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-743 <STI>
A:Cross-references: GB:M99436; NID:g307511; PIDN:AAA61193.1; PID:g307512
C:Genetics:
A:Gene: GDB:TLE2; ESG: ESG2
A:Cross-references: GDB:228048; OMIM:601041
A:Map position: 19p13.3-19p13.3
C:Superfamily: unassigned WD repeat proteins; WD repeat homology
C:Keywords: nucleus
F:499-532/Domain: WD repeat homology <WD1>
F:585-618/Domain: WD repeat homology <WD2>
F:667-700/Domain: WD repeat homology <WD3>
F:708-741/Domain: WD repeat homology <WD4>
Query Match 5.1%; Score 110.5; DB 2; Length 743;
Best Local Similarity 21.1%; Pred. No. 6.7;
Matches 87; Conservative 54; Mismatches 143; Indels 129; Gaps 20;
QY 14 KOLGALQASVLSLITENQGRKCFCAQNLMTQNTLPSVSHRSPPGNAVSVTGD 73
Db 101 QOVLOAVERAKQVTV-----GELNSLIGQQ--LQPLSHAPP----- 135
QY 74 CHLPTTEEFVGLVQSMKCDTVRIKGVLOQPTTAPLMTSEGNVTAEDTEAI---RAFV 129
Db 136 --VPL-----TPRAGLVGSGATG--LLALSGALAAQAALAAVKEDRAGV 177
QY 130 YAV-----AASAAEAHWRHLVLLSGIHEPIGSGNIIINTNKGGRSQNPALSPDQ 183
Db 178 EAEGSRVERAPRSASPSPPESLV---EERPSGPG---GGKQ-----RADEK 220
QY 184 SPSGNATTSVTRDNYHLITTEEFVWQSMKWHQSKSGSVPV-----RGPTQPCSES 238
Db 221 EFSGYEDEDKSDYNLVVDE-----QPSEPPSPATPCGVKVICPARRDLVDPASIA 276
QY 239 QLKESF-----VPTTPKENNKQEREDENWRLLPPVPAETPVPSVTEIETPL 288
Db 277 SLRSPLPRAKELIINDLPASTPASKSCDS-----SPQDASTPGPSSASHLCQAL 328
QY 289 QRIPRTATIA-GEPLGHCTFTTISPAFV-----HVLNKKRKQLELLREVEWPGRG--- 338
Db 329 KPAPSTDVSLRSLP-----TLSSPFTTSLGSHSTLN-----GDLSPSPSVYS 373
QY 339 HMAATCCCKLQVEGQDRMTSLAAAPVREAPPPPTGASSEPSVPALPGADPQRS 390
Db 374 LHLSPQSVSSVYGRSPVMAF-----ESHPLRGSSVSSSLPSIPGGKPAYS 420
Search completed: December 25, 2002, 01:14:13
Job time : 50 secs
```